https://ntrs.nasa.gov/search.jsp?R=19970015071 2020-06-16T02:04:46+00:00Z

NASA-CR-204195

FINAL
IN-53-CR
024595

FINAL ANNUAL REPORT FOR NAGW-4127

MODELING VISUAL, VESTIBULAR AND OCULOMOTOR INTERACTIONS IN SELF-MOTION ESTIMATION

DR. JOHN PERRONE
UNIVERSITY OF WAIKATO
PRIVATE BAG 3105
HAMILTON
NEW ZEALAND

Period covered: 5/31/95 - 6/31/96

2/4/97

John Perrone
Senior Lecturer (Associate Professor)

Summary.

A computational model of human self-motion perception has been developed in collaboration with Dr. Leland S. Stone at NASA Ames Research Center. The research included in the grant proposal sought to extend the utility of this model so that it could be used for explaining and predicting human performance in a greater variety of aerospace applications. This extension has been achieved along with physiological validation of the basic operation of the model.

Update of progress since Final Semi-Annual Report (5/31/95).

- Manuscript completed and submitted which reports the results of a large number of simulations of the model against existing physiological data from area MST of primate visual cortex (see Appendix 1).

- Two-dimensional motion sensors were developed with properties similar to those found in area MT (Middle Temporal) of visual cortex. These sensors enable digitized video image sequences to be used at input to our self-motion model instead of theoretical velocity vector fields. It therefore greatly expands the scope and validity of the model (see Appendix 2)..

- The development of a realistic two-dimensional sensor lead to a mechanism for incorporating eye-velocity information at the level of the MT units in our model. This is still being developed in conjunction with the image-based implementation of the self-motion model.

- The potential for incorporating higher level information (acceleration) has been demonstrated.

Publications/Conferences:

Perrone JA (1994) Simulating the speed and direction tuning of MT neurons using spatiotemporal tuned V1-neuron inputs. Invest Ophthal Vis Sci Suppl 35:2158.

Stone LS, Perrone JA (1994) A role for MST neurons in heading estimation. Soc. for Neurosci. Abstracts 20:772.

Perrone JA (1996) Generating acceleration sensitive motion sensors from sets of spatio-temporal filters. Invest Ophthal Vis Sci Suppl 37:S750

Stone LS, Perrone JA (1997) Human heading estimation during visually simulated curvilinear motion. Vision Res 37-573-590.

324.4 A ROLE FOR MST NEURONS IN HEADING ESTIMATION

L.S. Stone[1] and J.A. Perrone[2]

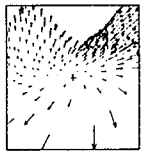
[1]Life Science Division, NASA Ames Research Center, U.S.A.
[2]University of Waikato, New Zealand

Introduction

- A template model that uses MT-like input elements can mimic human heading estimation under many conditions (Perrone, 1992; Perrone and Stone, 1994).
- The goal of this study is to compare the output elements of this model (heading detectors) with MST neurons.

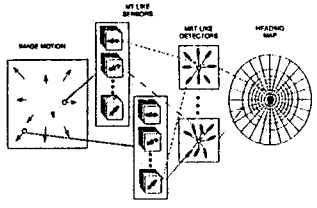
1 Background

A. The problem



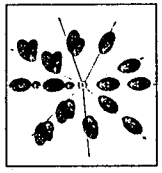
- Estimate direction of self-translation or heading (●) from a combined translation/rotation induced flow-field.

B. The template model



- Outputs of MT-like input sensors are combined by detectors.
- Maps of such detectors are used to estimate heading.

C. Heading detector



- Multiple MT-like sensors are fed in from each location in the visual field.

D. Design Principle

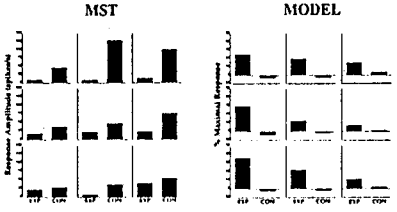
$$\begin{pmatrix}\dot{x}\\ \dot{y}\end{pmatrix}=\frac{v}{z}\begin{pmatrix}-1 & 0 & \tan\alpha_n\\ 0 & -1 & \tan\beta_n/\cos\alpha_n\end{pmatrix}\begin{pmatrix}\sin\alpha_n\cos\beta_n\\ \sin\beta_n\\ \cos\alpha_n\cos\beta_n\end{pmatrix}$$

$$+\frac{\omega_n}{\cos^2\alpha_n\sqrt{\sin^2\beta_n+\sin^2\alpha_n\cos^2\beta_n}}\begin{pmatrix}\sin\alpha_n\tan\beta_n & 1 & 0\\ \tan^2\beta_n+\cos^2\alpha_n & -\sin\alpha_n\tan\beta_n & 0\\ -\sin\alpha_n & 0 & 0\end{pmatrix}\begin{pmatrix}\sin\beta_n\\ -\sin\alpha_n\cos\beta_n\\ 0\end{pmatrix}$$

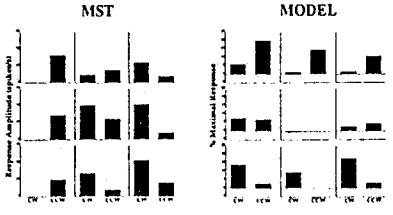
- Each detector is a template for a specific flow-field produced by combined translation/rotation self-motion.
- Rotation is assumed to result from gaze stabilization.

2 Position Invariance
(Duffy & Wurtz, 1991)

A. Radial Motion



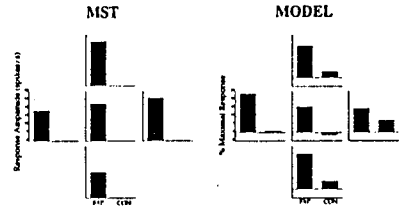
B. Roll Motion



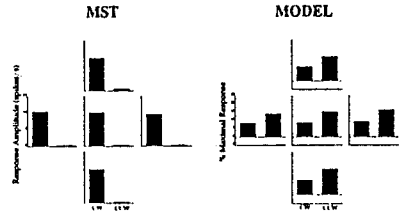
- Both MST neurons and model heading detectors show invariance and variance in Duffy paradigm.

3 Position Invariance
(Graziano et al., 1994)

A. Radial Motion

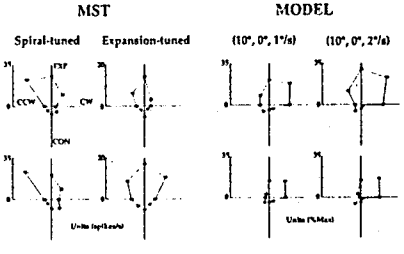


B. Roll Motion



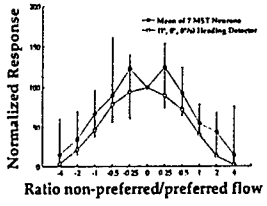
- Both MST neurons and model heading detectors show greater invariance in Graziano paradigm.

4 Spiral Invariance
(Graziano et al., 1994)



- Both MST neurons and model heading detectors show spiral invariance.

5 Template versus Decomposition
(Orban et al., 1992)



- Both MST neurons and model heading detectors act like templates rather than performing a decomposition of the flow-field.

Conclusions

- Our model heading detectors:
  - act like templates for specific instances of combined translation/rotation.
  - show the emergent properties of position and spiral invariance.
- Therefore neither position nor spiral invariance are incompatible with heading estimation.
- MST neurons:
  - act like templates.
  - show position and spiral invariance.
  - are therefore well-suited to support heading estimation.

References:

Duffy and Wurtz (1991) J. Neurophysiol., 65: 1329-1345.

Duffy and Wurtz (1991) J. Neurophysiol., 65: 1346-1359.

Graziano et al. (1994) J. Neurosci. 14: 54-67.

Orban et al. (1992) PNAS 89: 2595-2599.

Perrone (1992) JOSA A 9: 177-194.

Perrone and Stone (1994) Vision Research, 34: 2917-2938.

Investigative Ophthalmology & Visual Science,
March 15, 1994, Vol. 35, No. 4

4182 — 4:15

SIMULATING THE SPEED AND DIRECTION TUNING OF MT NEURONS USING SPATIOTEMPORAL TUNED V1-NEURON INPUTS.

((J. A. Perrone)) Psychology Dept., University of Waikato, New Zealand and NASA Ames Res.Ctr., Moffett Field, CA.

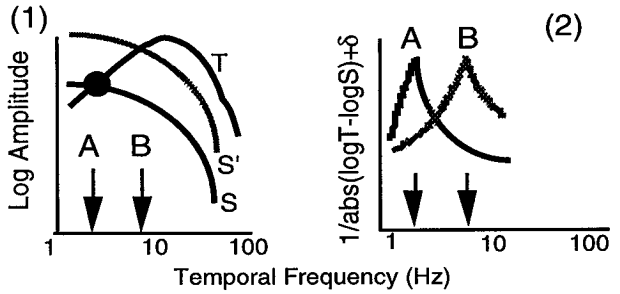
Purpose. Neurons in the primate middle temporal cortical area (MT) show characteristic speed and direction tuning useful for the extraction of self-motion and depth information from 2-D image motion (Perrone, JOSA 1992; Perrone & Stone, ECVP 1992). Neurons at the preceeding cortical level (V1) are tuned for particular spatiotemporal frequencies and can be modelled using linear filters (Watson & Ahumada, NASA TM 1983). However such filters are not velocity selective since their outputs are affected by factors such as the spatial frequency of the stimulus. Independence from such 'extraneous' stimulus features is desirable if the neuronal output is to be used for self-motion and depth extraction. **Method.** In order to construct a sensor, we combined the motion-energy outputs (Adelson & Bergen, JOSA 1985) from sets of linear spatiotemporal filters using a range of spatial frequencies but only two temporal-frequency channels (sustained and transient) as suggested by human psychophysics (Kulikowski & Tolhurst, J.Physiol. 1973). To set up a particular speed preference for the sensor as a whole, we adjusted the output ratio of the two temporal channels within each spatial-frequency band. Thus we were able to tune each band individually to the appropriate temporal frequency. **Results.** The sensor was tested with moving bars using a range of speeds and directions. The direction and speed tuning matched that of an "average" MT neuron. Using moving sine-wave grating inputs, we confirmed that the speed and direction tuning of the sensors is largely independent of the input spatial frequency. **Conclusion.** Motion-energy responses like those of directionally-selective V1 complex cells can be combined to create direction- and speed-tuned responses similar to those of MT neurons.

Supported by NASA RTOP #199-12-06-12-24 and NASA NCC2-307.

None

GENERATING ACCELERATION SENSITIVE MOTION SENSORS FROM SETS OF SPATIO-TEMPORAL FILTERS ((J. A. Perrone[1])) Psychology Dept., University of Waikato, New Zealand[1]

Purpose. Forward translation through the environment produces retinal image motion that often exhibits a large acceleration component. Acceleration information is useful for self-motion estimation and for the control of eye-movements (e.g., Pursuit. Krauzlis & Lisberger, Science, 1991). However such acceleration involves a continuous increase in the temporal frequency of the target and hence cannot be readily analyzed using simple spatio-temporal filters. We investigated methods for processing acceleration while still retaining the basic spatio-temporal filter architecture. **Methods.** Only two temporal-frequency channels (sustained = S and transient = T) are used. Adjusting the gain of the S channel alters the tf at which the two channel outputs are equal (see black dot in fig. 1). Subtraction of the log-transformed outputs of the T & S channels (plus inversion) produces an output tightly tuned to this tf (see A in fig.2). Thus speed tuning can be achieved by manipulating the gain of the S channel (Perrone, ARVO 1994). In order to construct a sensor tuned for acceleration (increasing tf) the gain of the S channel is altered to move S up to S'. This enables the increasing tf to be tracked (A to B), at least up to about 8 Hz. Summation of the outputs as the speed tuning moves from A to B will produce a large total 'acceleration' output if the target acceleration matches the rate specified by the A to B shift. **Results.** A wide range of speed and acceleration tunings were possible by applying the above mechanism across a small set of spatial frequency channels (4). **Conclusion.** Two broadly tuned temporal filters (supported by human psychophysics) are adequate for acceleration detection.



1. Supported by NASA grant NAGW-4127

None

APPENDIX 1: Abstract of paper submitted to Journal of Neuroscience.

EMULATING THE VISUAL RECEPTIVE FIELD PROPERTIES OF MST NEURONS WITH A TEMPLATE MODEL OF HEADING ESTIMATION

John A. Perrone[1], and Leland S. Stone[2]

[1]Department of Psychology
University of Waikato,
Hamilton, New Zealand

[2]Flight Management and Human Factors Division
NASA Ames Research Center
Moffett Field, CA, USA

ABSTRACT

We have previously proposed a computational neural-network model by which the complex patterns of retinal image motion generated during locomotion (optic flow) can be processed by specialized detectors acting as templates for specific instances of self-motion. The detectors in this template model respond to global optic flow by sampling image motion over a large portion of the visual field through networks of local motion sensors with properties similar to neurons found in the middle temporal (MT) area of primate extrastriate visual cortex. The model detectors were designed to extract self-translation (heading), self-rotation, as well as the scene layout (relative distances) ahead of a moving observer and are arranged in cortical-like heading maps to perform this function. Heading estimation from optic flow has been postulated by some to be implemented within the medial superior temporal (MST) area. Others have questioned whether MST neurons can fulfill this role because some of the receptive field properties appear inconsistent with those required for self-motion estimation. To resolve this issue, we systematically compared single-unit responses in MST with the outputs of model detectors under matched stimulus conditions. We found that most of the basic physiological properties of MST neurons can be explained by the template model. We conclude that MST neurons are well suited to support heading estimation and that the template model provides an explicit set of testable hypotheses which can guide future exploration of MST and adjacent areas within the primate superior temporal sulcus.

APPENDIX 2 (Manuscript of paper to be submitted to Nature).

PRECISE VISUAL SPEED ESTIMATION FROM BROADLY TUNED MOTION FILTERS.

John A. Perrone

Department of Psychology
The University of Waikato
Hamilton, New Zealand

Correspondence to:
Dr. J. A. Perrone
Department of Psychology
The University of Waikato
Private Bag 3105
Hamilton
New Zealand

e-mail: jpnz@waikato.ac.nz
Fax: 64 7 856 2158

Locomotion through the environment generates a pattern of image motion on the retina of our eyes [1]. The speed of the motion at a particular point on the retina is determined by a number of factors, such as our own speed, the retinal location of the motion, and the distance of the imaged object in the world [2,3]. Since most environments contain objects at a range of distances and because our visual field is large, the motion over the retina exhibits a great variety of speeds. This variation provides a rich source of information concerning the layout of the scene ahead if small speed differences can be detected over a wide range [2,3]. Humans possess such a speed discrimination ability [4] and neurones in the Middle Temporal (MT or V5) [5] area of primate visual cortex exhibit narrow speed tuning over a broad range of preferred speeds (1°/s - 512°/s) [6]. How this speed processing ability comes about is a long standing puzzle because the motion sensitive neurones found in the region prior to MT on the cortical visual pathway (area V1) are not speed tuned. I suggest a method by which this refinement in speed estimation between areas V1 and MT is achieved and demonstrate how the visual system is able to generate a wide range of speed tunings with very limited resources.

The consensus that has emerged after a number of theoretical, [7,8] psychophysical, [9] and physiological studies [10,11] is that visual motion is processed in area V1 of visual cortex by sets of spatio-temporal tuned neurones. These neurones respond maximally to particular combinations of spatial and temporal frequencies. Some of these V1 neurones (transient) are directionally selective for motion and have a biphasic temporal impulse response with band-pass temporal tuning [12,13] (Fig. 1a). Others (sustained) respond best to static features and are not directionally selective. These sustained neurones have a monophasic impulse response and low-pass temporal tuning[12,13] (Fig. 1b). Squaring of the outputs from appropriate pairs of these two types of motion filters produces a measure of the spatio-temporal 'energy' [14,15] at a particular retinal location. The amplitude response functions shown in Fig. 1b reflect the amount of energy generated for each temporal frequency.

Fig. 1 about here.

Both static and moving images are analysed by sets of filters of different spatial scales [16,17]. Each retinal location is represented by a number of different sized sustained and transient spatio-temporal filters (Fig. 2a). In the spatio-temporal frequency domain, an edge moving at speed v has a spectrum that lies along a line of slope $= -v$. [7,8] (Fig. 2a).

Fig. 2 about here

Motion of an edge at speed v deg/sec generates a temporal frequency equal to vu Hz in a spatial filter tuned to u cycles/deg. Within a particular spatial channel, changes to the speed (and hence the temporal frequency) will change the output of the transient filter in accordance with the temporal tuning function shown in Fig. 1b. However this temporal tuning is very broad compared to the speed tuning found in MT neurones [6] (see ahead to fig. 3b), and the output can be contaminated by changes to the spatial frequency and/or contrast of the input. An additional limitation of the spatio-temporal energy filters is that the set of possible preferred speed tunings is constrained by the small number of transient filters.

An ideal motion sensor tuned to speed v, must respond only to the spatial and temporal frequencies located along the line with slope = v. A single, transient spatio-temporal filter cannot meet this requirement on its own and information from a number of filters must be combined. This approach was used in an earlier model of image velocity estimation [18] but the filters did not have physiologically plausible temporal responses and differed somewhat from those shown in Fig. 1b. The popular gradient model of speed estimation [19, 20] also uses information from several filter types and incorporates division of transient channel outputs by sustained channel outputs. However, an important difference between the model to be described here and the gradient models is that the new motion filter generates an output that is speed tuned, not one that is linearly related to speed. Speed tuning is a well established property of MT neurones [6] whereas neurones that generate an output proportional to speed have never been found. Furthermore, mere speed tuning is perfectly adequate for the construction of template networks designed to model higher levels of motion processing such as the extraction of observer self-motion [21].

The solution I have developed to the speed tuning problem relies on a special combination rule for the sustained and transient filter energy outputs. An intuition for this process can be gained by noting that the two temporal tuning curves in Fig. 1b cross at one particular temporal frequency (see arrow). This is the point at which the outputs of the two filters in a particular spatial channel are equal. Consider a mechanism that produces a large output whenever the difference between the sustained and temporal filter outputs is zero (e.g., though a disinhibitory mechanism). For the temporal frequency corresponding to the position of the arrow in Fig. 1b, the output would be high. For temporal frequencies on either side, the absolute difference is not zero and the response would be less.

This feature becomes significant once it is realised that the 'cross-over' point can be manipulated by changing the gain of the sustained spatio-temporal filters. Note, for example, that a downward shift of the sustained curve in Fig.1b results in the cross-over point occurring at a lower specific temporal frequency. By using this gain mechanism, each spatial channel can be selectively 'tuned' to a particular temporal frequency (speed). In order to ensure that the maximum response occurs at the cross-over point we add a stage which sums the outputs from the transient and sustained filters. The mechanism for exploiting the cross-over point in each spatial channel can be formalised using the following equation that gives the filter output for a spatial channel, n:

$$F_n = \frac{LogT_n + Log(S_n \times G_n)}{\left|LogT_n - Log(S_n \times G_n)\right| + \delta} \qquad (1)$$

where T and S are the transient and sustained filter outputs and G is the gain used to weight the sustained filter output. For economy, the equation is intended to represent the functional aspects of the mechanism and not the underlying physiological implementation. The energy outputs of the transient and sustained filters are first passed through a compressive non-linearity (log) to compress their output range and to increase the sensitivity to low input contrast levels. The sustained filter outputs are weighted by the gain term (G) in order to set the cross-over point to the appropriate temporal frequency for the spatial channel, n. The sum of the two log-transformed and weighted outputs (logT and log(S x G)) is divided by a term which reaches a minimum when the transient and weighted sustained outputs are equal. The δ term prevents division by zero and sets the speed tuning bandwidth (full width at half-height) of the filter. Such an equation could effectively be implemented biologically via a disinhibition mechanism yielding true narrowly-tuned speed tuning within a single spatial channel as opposed to proportional speed outputs in spatial channels [22].

While moving us a step closer to the ideal filter, the mechanism outlined above only deals with the vertical temporal frequency axis. Because the spectrum of a moving edge is oriented relative to the horizontal spatial frequency axis (see Fig. 2a), the filter must be oriented in spatio-temporal frequency space. This orientation can be achieved by the judicious choice of the spatio-temporal filter properties.

I discovered that if two basic conditions exist in each spatial channel, then the application of the model described in Eq. 1 will produce the required oriented filter .

The first condition is that the peak spatial frequency of the transient filter within a linked pair be shifted slightly towards lower frequencies (~2.5%) relative to the sustained filter. The second condition is that the spatial bandwidth of the transient filter be slightly larger (~2%) than the bandwidth of the sustained filter. When these two conditions exist and the gain term for the channel is adjusted appropriately, a filter results that is close to our ideal and suited to the line spectrum generated by moving edges in the scene (Fig. 2b).

Figure 2b shows the situation in which all the spatial channels are tuned to the same speed (-4°/sec). Other configurations are possible but this arrangement will be used to demonstrate the new motion sensor. A combination rule is required if the information from all of the spatial channels is to be used to give an overall speed estimate. This is a general problem in image analysis and not specific to motion estimation [23]. I have adopted a scheme in which the maximum output across the four spatial channels ($Max[F_n]$, $n = 1,4$) is used as the final output of the combined motion sensor. Allowance must be made for the different sizes of the spatial filters in each channel and their different spatial sampling rates. In the simulations that follow, only the outputs from filters centred on one spatial location are considered. Solutions for the more general case of two-dimensional distributions of filters have been developed and will be presented in a future publication.

Figure 3 about here

With the mechanism outlined in the new model, a wide range of speed tunings can be achieved by simply changing the gain term within each spatial channel. Each channel can be tuned to any temporal frequency in the range from 0 to about 8 Hz. Within these upper and lower bounds, a huge variety of speed tunings can be set up using the same minimal channel architecture. The speed tuning for a number of model sensors tuned to a range of image speeds is shown in Fig. 3 along with reproduced data from MT neurones [6]. The model sensors exhibit the same peaked tuning functions seen in the physiological data, with their response falling to approximately 50% with a doubling or halving of the preferred speed of the sensor. The use of sampled digital imagery prevented the very highest speed tuning (256°/s) being simulated. Other than this minor limitation, the model is able to reproduce the speed tuning patterns found in MT neurone responses using well documented properties of neurones found in area V1 of primary visual cortex. It therefore offers an explanation of how the V1-MT speed tuning transformation occurs.

There already exists indirect evidence for the weighting mechanism proposed in the model. In order to tune each spatial channel to the same speed v, the equal-output cross-over points for the transient and sustained filter outputs need to occur at a low temporal frequency in low spatial frequency channels and at a high temporal frequency in high spatial frequency channels. For this to occur, the sustained filter outputs must be reduced by a large amount in the low spatial frequency channels and increased in the high spatial frequency channels. If such a weighting pattern existed we would expect to find that the transient filters dominate at low spatial frequencies and the sustained filters dominate at high spatial frequencies. One would also predect a systematic change in the relative sensitivities of the transient and sustained filters as their preferred spatial frequency changes. These results have often been found in a variety of psychophysical experiments. [24, 25]

I have demonstrated that despite what at first appears to be minimal and inadequate resources in area V1 (just two broadly tuned temporal channels and a limited number of spatial channels), it is possible to derive precise speed tuning over a wide range of preferred speeds. The proposed mechanism could form the basis of a general process by which biological systems obtain very fine perceptual discriminations from the broadly tuned filters common to many of the senses.

References:

1. Gibson, J.J. *The Perception of the Visual World.* (Houton Mifflin, Boston, 1950)
2. Koenderink, J.J. & van Doorn, A.J. Opt. Acta. **22,** 773-791 (1975).
3. Longuet-Higgins, H.C. & Pradzny, K. Proc. R. Soc. Lond. **B208**, 385-397 (1981).
4. McKee, S.P. *Vision Res.* **21**, 491-500 (1981).
5. Zeki, S.M. *J. Physiol. (Lond).* **236**, 549-573 (1974).
6. Maunsell, J.H.R. & Van Essen, D.C. *J. Neurophysiol.* **49**, 1127-1147 (1983).
7. Fahle, M. & Poggio, T. *Proc. R. Soc. Lond.* **B213,** 415-477 (1981).
8. Watson A.B. & Ahumada, A.J. in *Motion: Perception and representation* (ed. Tsotsos, J.K) (New York: Association for Computing Machinery). 1-10 (1983).
9. Burr, D., Ross, J. & Morrone, M.C. *Proc. R. Soc. Lond.* B**247**, 249-265 (1986).
10. Emerson, R.C., Citron, M.C., Vaughn, W.J. & Klein, S.A. *J. Neurophysiol.* **58,** 33-65 (1987).
11. McLean, J. & Palmer, L.A. *Vision Res.* **29**, 675-679 (1989).
12. Kulikowski, J.J. & Tolhurst, D.J. *J. Physiol.* **232**, 149-162 (1973).
13. Hawken, M.J., Shapley, R.M. & Grosof, D.H. *Vis.Neurosci.* **3**, 477- 492 (1996)
14. Adelson, E.H. & Bergen, J.R. *J. Opt. Soc. Am.* **2**, 284-299 (1985).
15. Emerson, R.C., Bergen, J.R. & Adelson, E.H. *Vision Res.* 32, 203-218 (1992).
16 Campbell, F.W. & Robson, J.G. *J. Physiol. Lond.* **197**, 551-566 (1969).
17. Adelson, E.H. & Movshon, J.A. *Nature* **300**, 523-525 (1982).
18. Heeger, D.J. J.Opt. Soc. Am. A. **4**, 1455-1471 (1987).
19. Marr, D. & Ullman, S. *Proc. R. Soc. Lond.* B**211**, 151-180 (1981).
20. Harris, M.G. *Vision Res.* **26**, 1281-1287 (1986)
21. Perrone, J.A. & Stone, L.S. *Vision Res.* **34**, 2917-2938 (1994).
22. Adelson, E.H. & Bergen, J.R. in *Workshop on motion: Representation and analysis* (IEEE Computer Society Press) 151-155 (1986).
23. Braddick, O. *Nature*, **320**, 680-681 (1986).
24. Robson, J.G. *J. Opt. Soc. Am.* . **56**, 1141-1142 (1966).
25. Anderson , S.J. & Burr , D.C. *Vision Res.* **25**, 1147-1154 (1985).
26. Watson, A.B. in *Handbook of Perception and Human Performance* (eds. Boff, K., Kaufman, L. & Thomas, J.) (Wiley, New York, 1986).
27. Watson, A.B. & Ahumada, A.J. *J. Opt. Soc. Am.* **2**, 322-342 (1985).

Acknowledgement: Thanks to Leland Stone and Bernard Guerin for their helpful comments.

Figure Captions.

Figure. 1.
(a) Temporal impulse responses of the transient and sustained spatio-temporal filters used in the model. The temporal impulse response extends over a 160 msec period. The sustained function is monophasic and favours integration of static features. The transient function is diphasic, responding best to temporally modulating (moving) stimuli. (b) Amplitude response functions used to model the temporal frequency tuning of the transient and sustained spatio-temporal filters. They are based on equations derived to fit human psychophysical temporal sensitivity data [26] . The arrow indicates the temporal frequency at which the output of the two filter types is equal.

Figure 2.
(a) Frequency domain representation of the spatio-temporal filters used to generate the speed tuned motion filter. Only the upper right quadrant of frequency space is depicted. Four spatial channels are used although the exact number is not critical to the discussion. A vertical slice through one of the spatial channels would reveal the profiles depicted in the amplitude response curves of Fig. 1 (b) . The model spatio-temporal filters are constructed using the design of Watson & Ahumada. [8, 27] The transient channels were set to 4, 8, 16, and 32 cycles/width (width = 32° = 256 pixels in model simulations). An edge moving from right to left at a particular speed -v, generates a spectrum of slope = v. The aim is to construct a filter that responds selectively to only one slope. Image motion in directions other than 180° has the effect of moving the spectrum in a plane passing through the line shown in Fig. 1b, and is easily dealt with by the inclusion of spatio-temporal filters tuned to different orientations.

b) New speed tuned filter tuned to -4°/sec. The sustained spatial filter bandwidths were first set to 1 octave and the transient centre spatial frequencies set as specified in a). The gain term (G_n) and filter parameters required in each channel were then determined from the amplitude response functions of the spatial and temporal channels using a search algorithm which minimised the 'spread' of the filter output around a line of slope = 4.0. The log- frequency bandwidth of the transient filters was 1.02 octaves and the sustained centre spatial frequencies were a factor of 1.025 higher than those of the transient filters. The delta term in Eq. 1 determines the speed tuning bandwidth of the filter and it was set to 0.4. In this example, the different spatial channels are all tuned to the same speed and the maximum output across the 4

channels was used as the total sensor output. This type of sensor is largely independent of the spatial frequency content of the moving image.

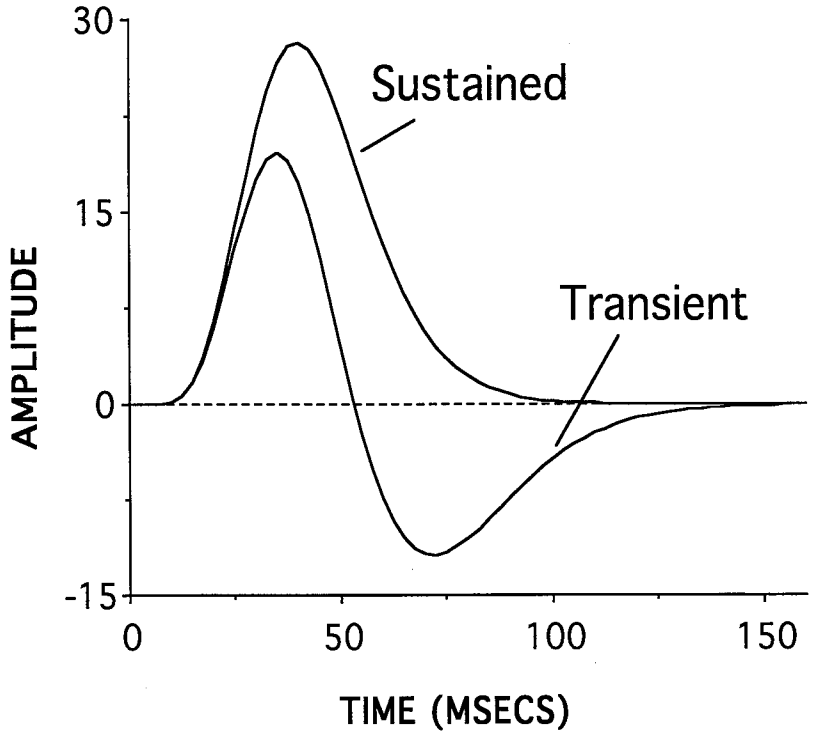
Figure 3.

a) Speed tuning results from a simulation using digital image sequences (256 pixels wide x 8 frames). It was assumed that 256 pixels map onto 32 degrees of visual field and that the frames were sampled at a 64Hz rate. Thus 1°/s was equivalent to 0.125 pixel/frame. Digital implementation of the sensor requires a sampled version of the temporal filters which introduces some errors. For this reason, the gains were set empirically by determining the sustained/transient output ratios for a range of edge speeds. Model sensors tuned to 4°/s (open squares), 16°/s (filled circles), 64°/s (open triangles) and 128°/s (open circles) were tested using edge speeds ranging from 1°/s to 256°/s in octave steps. Speeds higher than 256°/s could not be tested because of the image size limits. The model filters were always located at the midpoint of the edge's travel and the maximum output across the four spatial channels was used as the final output of the sensor. The outputs for each sensor were normalised with respect to the peak tuning response and these values are plotted in the graph.

b) Reproduced data from Maunsell & Van Essen [6] who tested the speed tuning of neurones in area MT of primate visual cortex over a wide range of stimulus speeds. They found neurones with preferred speed tunings covering a broad range, but the majority of the cells in their sample were tuned to approximately 32°/s. The responses have been normalised to the maximum output for the cell. The open circles and dashed lines represent responses that were below the spontaneous firing rate for the cell.

Fig. 1

(a)


30
15
0
-15
AMPLITUDE
Sustained
Transient
0
50
100
150
TIME (MSECS)

(b)

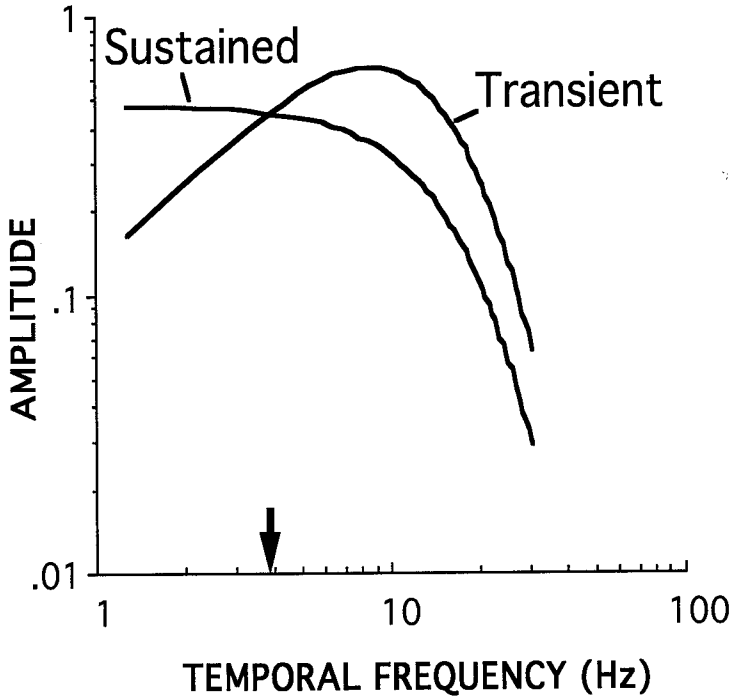

1
.1
.01
AMPLITUDE
Sustained
Transient
1
10
100
TEMPORAL FREQUENCY (Hz)

Fig. 2

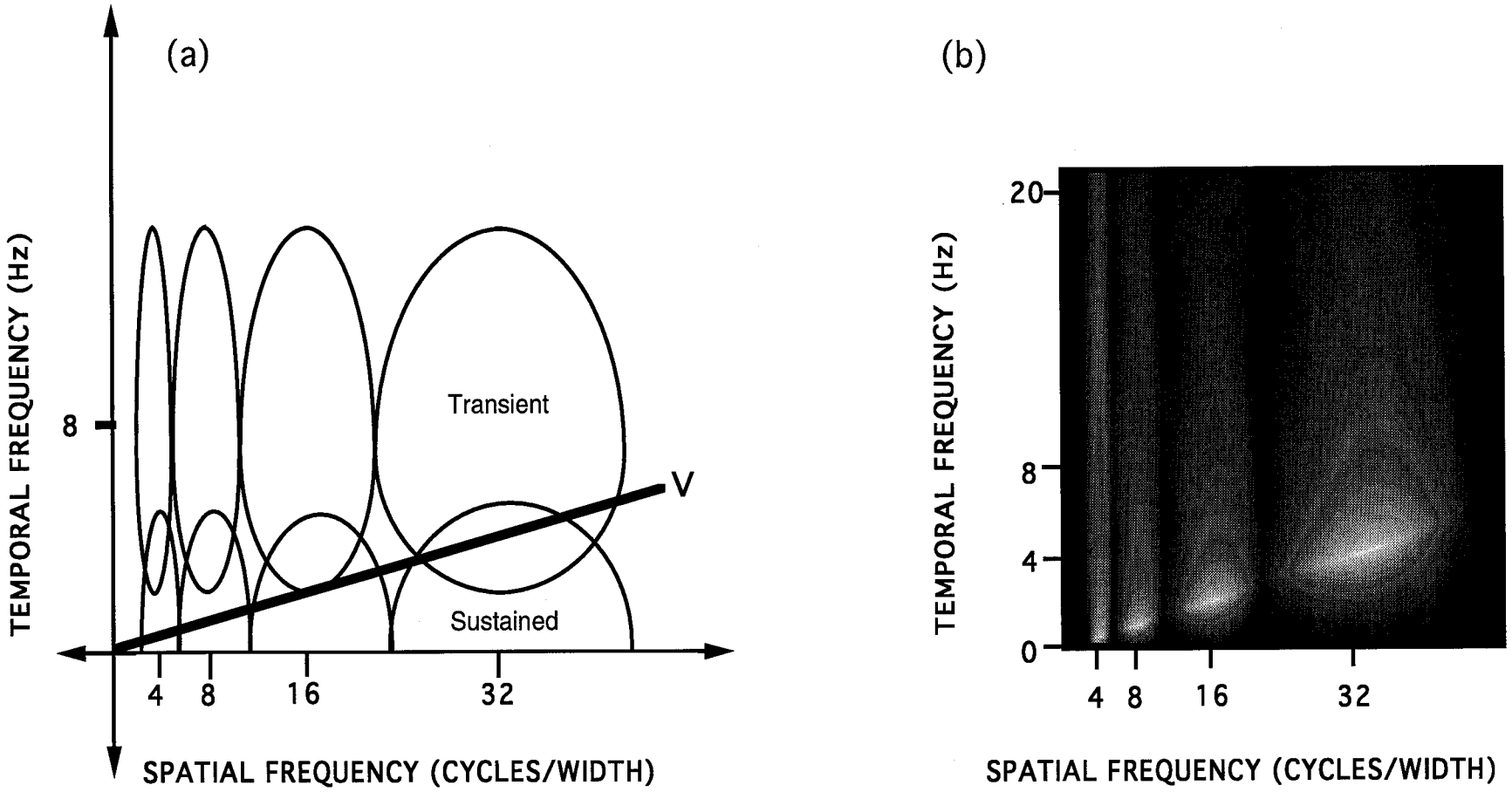
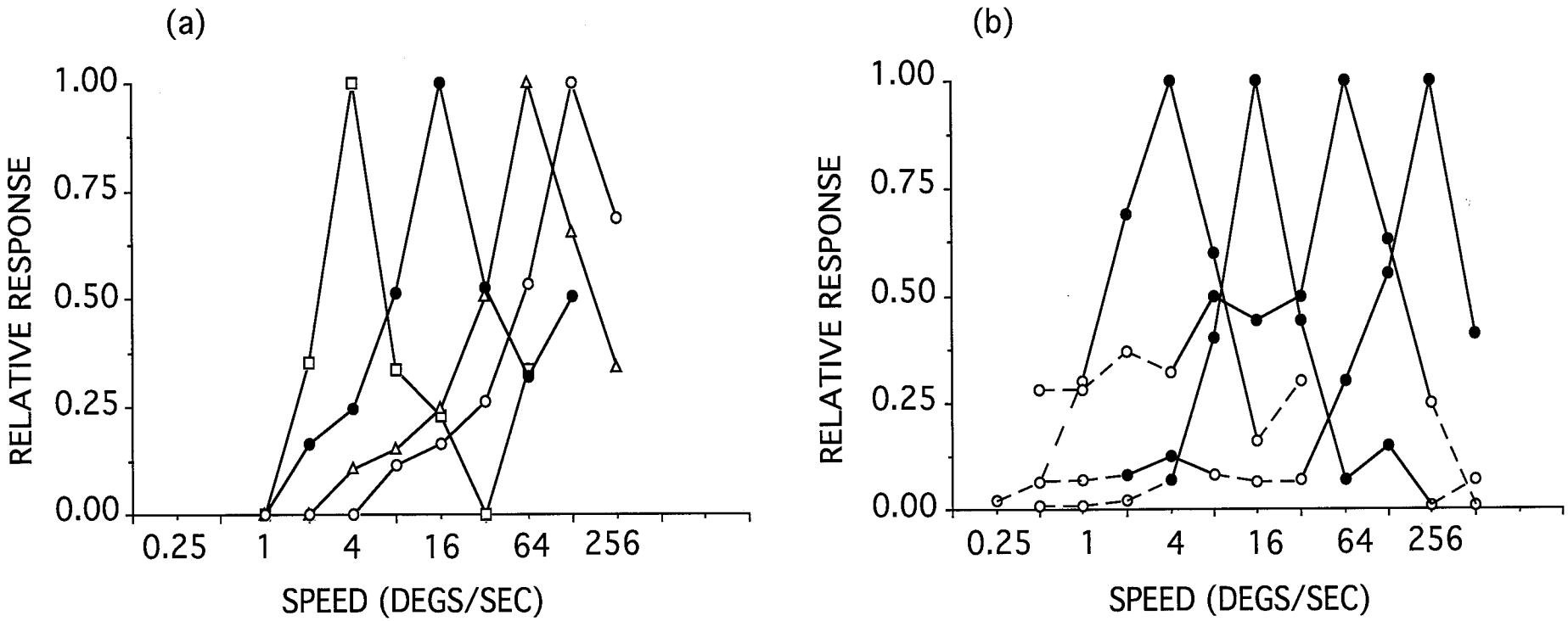

(a)
TEMPORAL FREQUENCY (Hz)
8
Transient
V
Sustained
4
8
16
32
SPATIAL FREQUENCY (CYCLES/WIDTH)
(b)
TEMPORAL FREQUENCY (Hz)
20
8
4
0
4
8
16
32
SPATIAL FREQUENCY (CYCLES/WIDTH)

Fig. 3


(a)
1.00
0.75
0.50
0.25
0.00
RELATIVE RESPONSE
0.25
1
4
16
64
256
SPEED (DEGS/SEC)
(b)
1.00
0.75
0.50
0.25
0.00
RELATIVE RESPONSE
0.25
1
4
16
64
256
SPEED (DEGS/SEC)